EXHIBIT B
(Part 2 of 3)

sequence blocks for the upstream (or downstream) oligos that are contained in the filtered address sets for a given address block, the following rule is adhered to: a sequence block only occurs in the upstream subset (or in the downstream subset) of an address block, if every oligo that is contained in that address block occurs in the upstream (or in the downstream) subset of every filtered address set that pertains to that address block. For example, sequence block "CATG" occurs in the upstream subset of address block "TACCTTG" because oligos "CAT" and "ATG" occur in the upstream subset of address oligos "TAC", "ACC", "CCT", "CTT", and "TTG".

Often, a sequence block does not occur in its own upstream or downstream subset. For example, sequence block "CATG" does not occur in the upstream or downstream subset of its own block set (i.e., in block set "CATG"), because oligo "ATG" is not present in the upstream subset of address set "CAT" and oligo "CAT" is not present in the downstream subset of address set "ATG". When a sequence block does not occur in its own upstream or downstream subset, this indicates that that sequence block occurs only once in the nucleotide sequence of that strand. However, a sequence block may occur in both the upstream subset and in the downstream subset of its own block set. For example, sequence block "TGGTA" occurs in both the upstream subset and in the downstream subset of block set "TGGTA". When a sequence block does occur in its own upstream and downstream subsets, it indicates that the sequence block may, but not must, occur more than once in the sequence. The presence of more than one parental strand in the original mixture can introduce additional oligos into the filtered upstream and downstream subsets that can cause a block that actually occurs only once in a sequence to appear in both the upstream and downstream subsets of its own block set. However, further analysis of the data determines the multiplicity of each block in the strand (as described below), thus resolving these uncertainties. For convenience, block sets that pertain to blocks that definitely occur only once in the sequence are listed together. For example, in Figure 10d, block set "CATG" and block set "TACCTTG" are listed together.

Fourth, the position of each sequence block relative to the other sequence blocks is determined. An examination of the block sets that pertain to unique blocks (that definitely occur only once in the sequence of the strand) indicates their relative positions. For example, in Figure 10d, block set "CATG" indicates that unique sequence block "TACCTTG" occurs downstream of unique sequence block "CATG". This is confirmed by block set "TACCTTG", in which unique sequence block "CATG" occurs upstream of unique sequence block "TACCTTG". The relative position of the two unique sequence blocks is indicated in Figure 10e, where the top line to the left of the arrow shows "CATG" upstream (to the left) of "TACCTTG". The relative position of the sequence blocks that can potentially occur more than once in the nucleotide sequence of the strand is determined from their presence or absence in the upstream and downstream subsets of other sequence blocks. For example, sequence block "TAA" occurs in the downstream subset of block set "CATG" (and does not occur in the upstream subset of block set "CATG"). Furthermore, sequence block "TAA" also occurs in the downstream subset of block set "TACCTTG" (and not in its upstream subset). Therefore, sequence block "TAA" must occur downstream of both unique sequence blocks "CATG" and "TACCTTG". This is indicated in Figure 10e, where the bottom line to the left of the arrow shows "TAA" as occurring downstream of "CATG" and "TACCTTG". Furthermore, sequence block "TGGTA" occurs only in the downstream subset of block set "CATG". Therefore, it must occur downstream of "CATG" in the sequence. On the other hand, sequence block "TGGTA" occurs in both the upstream and downstream subsets of block set "TACCTTG". This indicates that "TGGTA" can potentially occur in the sequence at positions both upstream and downstream of unique sequence block "TACCTTG". Finally, "TGGTA" only occurs upstream of "TAA". This is indicated in Figure 10e, where the bottom line to the left of the arrow contains a bracket that shows the range of positions at which "TGGTA" can occur, relative to the positions of the other sequence blocks. At this point in the analysis, the diagram to the left of the arrow in Figure 9c contains all the information obtained that pertains to strand set A.

Finally, the sequence of the strand is ascertained by taking into account both the relative position of the sequence blocks, as shown in the diagram to the left of the arrow in Figure 10e, and the identity of the sequences at the ends of the sequence blocks. The object of this last step is to assemble the blocks into the final sequence. Four rules are followed: (a) each of the blocks must be used at least once; (b) the blocks must be assembled into a single sequence; (c) the ends of blocks that are to be joined must maximally overlap each other (i.e., if the surveyed oligos are *n* nucleotides in length, then two blocks maximally overlap each other if they share a terminal sub-sequence that is *n*-1 nucleotides in length); and (d) the order of the blocks must be consistent with their positions relative to one another, as ascertained from the block sets. For example, in Figure 10e, "CATG" is upstream of "TACCTTG". "CATG" cannot be joined directly to "TACCTTG", since these two sequence blocks do not possess maximally overlapping terminal sequences (two nucleotides in length). However, an examination of the permissible positions at which other sequence blocks can occur indicates that "TGGTA" can occur in the gap between "CATG" and "TACCTTG". The ends of these sequence blocks are then examined to see whether the gap can be bridged. "CATG" can be joined to "TGGTA" by maximally overlapping their shared terminal sub-sequence "TG". Furthermore "TGGTA" can be joined to "TACCTTG" by maximally overlapping their shared terminal sub-sequence "TA". Similarly, the gap that occurs downstream of "TACCTTG" can potentially be filled by both "TAA" and "TGGTA". "TAA" must be used, because it was not used at any other location. However, "TACCTTG" cannot be directly joined to "TAA". The solution is to join "TACCTTG" to "TGGTA", and then to join "TGGTA" to "TAA". Thus, the sequence of strand A (which is shown in Figure 10f) is unambiguously assembled by utilizing sequence block "TGGTA" twice (as summarized in the diagram to the right of the arrow in Figure 10e).

The same procedure is followed to determine the sequence of strand B (see Figure 11). In this example, there are three sequence blocks that do not occur in their own upstream or downstream subsets, and they therefore definitely occur only once

in the sequence of strand B (namely, sequence blocks "CTTG", "GTCC", and "TACC"). An examination of block set "GTCC" shows that "GTCC" occurs upstream of "CTTG" and "TACC". However, an examination of block set "CTTG" and an examination of block set "TACC" indicates that sequence blocks "CTTG" and "TACC" can both occur upstream and downstream of each other, which appears to conflict with the observation that these sequence blocks only occur once in the sequence of strand B. There is actually no conflict. Each of these sequence blocks does indeed occur only once. It is just that their positions, relative to one another, in strand B are obscured by the presence of conflicting information from the relative positions of oligos that occur in strand A. This ambiguity (indicated by the identical positions of sequence blocks "CTTG" and "TACC" in the diagram to the left of the arrow in Figure 11e) is resolved by the remainder of the information. The positions of those sequence blocks that can potentially occur more than once in the sequence of strand B is determined from other block sets. First, the block sets of the sequence blocks that definitely occur only once in the sequence (namely, block sets "CTTG", "GTCC", and "TACC") are consulted. The range of positions at which these other sequence blocks can occur (relative to the positions of other blocks) is indicated in the diagram to the left side of the arrow in Figure 11e.

The assembly of the nucleotide sequence of Strand B proceeds as follows: "ATG" is upstream of all other blocks. The uniquely occurring block immediately downstream of "ATG" is "GTCC". "ATG" and "GTCC" cannot be directly joined. However, "ATG" can be directly joined to "TGGT", so the correct order is to join "ATG" to "TGGC", and then to join "TGGC" to "GTCC". Neither "CTTG" nor "TACC" can be directly joined to "GTCC". Three different sequence blocks can be used to bridge this gap (namely, "CCT", "GTA", and "TGGT"). The only combination of these three sequence blocks that can fill this gap is "CCT" alone, which bridges the gap between "GTCC" and "CTTG". This resolves the ambiguity as to the relative positions of "CTTG" and "TACC". "CTTG" is therefore upstream of "TACC". "CTTG" cannot be directly joined to "TACC". Again, there are three different sequence blocks that can be used

to fill this gap (namely, "CCT", "GTA", and "TGGT"). The only combination of these three sequence blocks that can fill this gap is "TGGT" and "GTA" (i.e.,"GTTG" is joined to "TGGT", "TGGT" is joined to "GTA", and "GTA" is joined to "TACC"). And finally, "CTA", which occurs upstream of all other blocks, must be included in the sequence. However, "TACC" cannot be directly joined to "CTA". There are three different sequence blocks that can be used to fill this gap (namely, "CCT", "GTA", and "TGGT"). The only combination of these three sequence blocks that can fill this gap is "CCT" alone. Thus, the assembly of the sequence of Strand B from its sequence blocks is completed. Note that some sequence blocks that could potentially occur in the sequence more than once, actually occur only once (e.g., "GTA"), while others actually occur more than once (e.g., "CCT").

Using the methods of this invention, the entire sequence of strand B is unambiguously determined, despite the fact that some oligos occur more than once in its sequence, despite the fact that more than one sequence block can be assembled from the oligos that occur in the strand, despite the fact that the multiplicity of occurrence of each oligo is not determined during surveying, despite the fact that the strand is analyzed in a mixture of strands, and despite the fact that the other strand in the mixture possesses many of the same oligos.

5. Uses of sectioned oligonucleotide arrays for manipulating nucleic acids --

In the examples described below, it is assumed that the sequences of the nucleic acids to be manipulated have already been established. It is not necessary, in these manipulations, that the sample be distributed across the entire array. Instead, a sample can be delivered directly to the well in the array where a particular oligo (or a particular strand) is immobilized. The arrays enable a large number of specifically directed manipulations of nucleic acids to be carried out.

5.1. Cleavable primers --

Amplification of strands and partials following separation (or generation) on a sectioned array requires that their ends be provided with priming regions. The priming regions can be undesirable in subsequent use, such as the making of recombinants or site-directed mutants. For some uses it is desirable to substitute new priming regions for the old. For those uses, the primers used for amplification must first be removed from the 5' ends.

Where the junction of the primer and the strand is contained within a unique restriction site, the primer can be removed by treating a double-stranded version of the strand with a corresponding restriction endonuclease. However, restriction sites will often not be present at the junctions. A solution to this problem is to make the primer (or even only the junction nucleotide in the primer) chemically different from the rest of the strand. The primer in these examples resides at the strand's 5' terminus.

5.1.1. Cleavage of primers by alkaline hydrolysis or by ribonuclease digestion --

This method is suitable for removal of oligoribonucleotide primers, or mixed RNA/DNA primers whose 3' terminal nucleotide (which becomes a junction nucleotide upon primer extension) is a ribonucleotide. Such primers are incorporated at the 5' end of DNA strands or partials during amplification.

Alkaline hydrolysis cleaves a phosphodiester bond that is on the 3' side of a ribonucleotide, and leaves intact a phosphodiester bond that is on the 3' side of a deoxyribonucleotide. After alkaline hydrolysis, the pH of the reaction mixture is returned to a neutral value by the addition of acid, and the sample can be used without purification. Primers containing a riboadenylate or a riboguanylate residue at their 3' end can effectively be removed from a DNA strand or partial by treatment with T_2 ribonuclease. After treatment, the sample is heated to 100°C to inactivate the ribonuclease, and can be used without purification. In both these cases, the released 5' terminus of

the strand (or partial) is left dephosphorylated. Therefore, if the strand obtained is subsequently used for ligation, it should be phosphorylated by incubation with polynucleotide kinase.

5.1.2. Cleavage of primers from DNA strands (or partials) synthesized from phosphorothioate nucleotide precursors --

In this method, oligodeoxynucleotide or oligoribonucleotide primers are synthesized from natural nucleotides, but strand amplification is carried out in the presence of only α-phosphorothioate nucleotide precursors. Subsequent digestion of the synthesized strands with a 5'-3' exonuclease, such as calf spleen 5'-3' exonuclease, results in the elimination of all primer nucleotides except the original 3'-terminal (junction) nucleotide of the primer, with the released 5'-terminal group of a strand or partial being unphosphorylated. The junction nucleotide is not removed, because it is joined to the rest of the strand by a phosphorothioate diester bond. Therefore, the strand obtained has an extra nucleotide at its 5' end. This does not present a problem when the presence of the former junction nucleotide at the 5' end of the strand is compatible with the subsequent use of the strand. The presence of the extra nucleotide can also be useful for site-directed mutagenesis.

If the primer-deprived strand so obtained is to be ligated, the use of spleen exonuclease, which leaves 5'-hydroxyl groups, must be then followed by phosphorylation with polynucleotide kinase. Therefore, where the strand is to be ligated, the use of bacteriophage lambda or bacteriophage T7 5'-3' exonuclease is preferable over spleen exonuclease, since they leave 5'-phosphoryl groups at the site of cleavage.

5.2. Generation of recombinant nucleic acids --

In the method described below, two nucleic acid strands are ligated in one round of ligation. It is possible to keep repeating the process any desired number of times to ligate the desired number of strands.

In this example, a sectioned array contains immobilized oligos that consist of two portions, one complementary to the 3'-

terminal sequence of one of the moieties to be ligated, and the other complementary to the 5'-terminal sequence of the other moiety to be ligated. The immobilized oligos can have either free 3' or 5' ends. The relevant termini of the moieties to be ligated should be deprived of priming regions, but priming regions (preferably different) should be preserved at the opposite termini to allow amplification of the recombinants. After hybridization in an appropriate well, the two nucleic acid strands are ligated to each other utilizing DNA ligase. Unligated strands are then washed away. Only ligated strands possess two terminal priming regions required for PCR. The strands that are to be ligated can be used in a mixture with other strands, provided that no other strands have with the same oligos at the termini deprived of priming regions.

Many different strands can be ligated to one particular strand (or partial), to produce many recombinant variations of one gene. In that case, one portion of the splint, i.e., the immobilized oligo is a constant segment, and the other portion is a variable segment, i.e., a binary array is used. The constant segment binds to the strand to be included in every recombinant, and the variable segment binds to the end of a strand to be fused with the invariant strand.

5.3. Site-directed mutagenesis --

The ability to prepare any partial of a strand according to the invention provides the opportunity to make nucleotide substitutions, deletions and insertions at any chosen position within a nucleic acid. Moreover, the use of sectioned arrays makes it possible to perform site-directed mutagenesis at a number of positions (even at all positions) at once, and in a particular embodiment, to determine, within individual wells of the array, properties of the encoded mutant proteins.

Mutations are introduced into a strand by first preparing partials having variable ends that correspond to the segment to be mutated, that segment preceding the location of the intended mutation. Then mutagenic nucleotides or oligos are introduced into the variable ends. The mutated partials are then extended

the length of the full sized strand using the complementary copy of the original non-mutated strand as a template.

In this method, complements of partials (i.e., strands whose 5' termini are variable and 3' termini are fixed) are used. Their 5'-terminal priming regions are removed and then phosphorylated by incubation with polynucleotide kinase, and the partials are then ligated by incubation with RNA ligase to the free 3' hydroxyls of oligoribonucleotides immobilized on a 3' sectioned ordinary array. The sequence of the immobilized oligo to which a partial is ligated is identical to the oligo segment that occurs in the original (full-length) strand immediately adjacent to the end of the partial, except for one (or a few) nucleotide difference(s) that corresponds to mutation(s) to be introduced.

The nucleotide differences are preferably located at the 3' terminus of the immobilized oligo, and can correspond to a nucleotide substitution, insertion, or deletion. A deletion can be of any size. For a large insertion, the ligated partial, or the immobilized oligo, can first be fused to a nucleic acid containing all or part of the sequence to be inserted.

After washing away material not covalently bound, the immobilized strand is linearly copied, taking advantage of the priming region at its (fixed) 3' end. The copies correspond to partials that have been extended by the oligos containing the mutation(s). The copies are annealed to their complementary full-length strands, and their 3' termini extended by incubation with DNA polymerase, using the parental strand as a template. Finally, the extended mutant strands are amplified by PCR. It is important that the primers utilized for amplification of a partial used for mutagenesis be different from the primers used to amplify the original (non-mutant) full-length strand. This assures that only mutant strands are amplified.

We claim:

1. A binary oligonucleotide array comprising an array of predetermined areas on a surface of a solid support, each area having therein, covalently linked to said surface, multiple copies of a binary oligonucleotide of a predetermined sequence, said binary oligonucleotide consisting of a constant nucleotide sequence adjacent to a variable nucleotide sequence, wherein the constant nucleotide sequence is the same for all oligonucleotides in the array.

2. A binary array according to claim 1 wherein the binary oligonucleotides consist of deoxyribonucleotides.

3. A binary array according to claim 1 wherein the binary oligonucleotides consist of ribonucleotides.

4. A binary array according to claim 1 wherein one or more of nucleotides of the binary oligonucleotides are modified.

5. A binary array according to claim 1 wherein one or more of the nucleotides of the binary oligonucleotides are non-standard.

6. A binary array according to claim 1 wherein the binary oligonucleotides are mixed.

7. A comprehensive binary array according to claim 1

8. A comprehensive binary array according to claim 7 wherein the binary oligonucleotides in each area have variable sequences of the same length.

9. A 3' binary array according to claim 1.

10. A 5' binary array according to claim 1.

11. A 3' binary array according to claim 9, wherein each covalently linked binary oligonucleotide has its constant sequence adjacent to the 5' end of its variable sequence.

12. A 5' binary array according to claim 10, wherein each covalently linked binary oligonucleotide has its constant sequence adjacent to the 3' end of its variable sequence.

13. A binary array according to claim 2 wherein all or part of the constant nucleotide sequence is complementary to a predetermined restriction recognition sequence.

14. A binary array according to claim 1 having an oligonucleotide hybridized to all or part of the constant sequence which is ligatable to the terminus of an adjacent nucleic acid hybridized to the oligonucleotide.

15. In an oligonucleotide array having variable-sequence oligonucleotides immobilized in a predetermined pattern of areas on a solid support, the improvement comprising including in said oligonucleotides a constant sequence of predetermined length.

16. A sectioned binary array according to claim 1.

17. A comprehensive sectioned binary array according to claim 16.

18. A 3' binary oligonucleotide array according to claim 17, wherein each covalently linked binary oligonucleotide has its variable sequence adjacent to the 5' end of its constant sequence.

19. A 5' binary oligonucleotide array according to claim 17, wherein each covalently linked binary oligonucleotide has its variable sequence adjacent to the 3' end of its constant sequence.

20. A binary oligonucleotide array according to claim 1, wherein said constant nucleotide sequence comprises one or more functional sequences selected from the group consisting of a nucleic acid polymerase priming region, an RNA polymerase promoter region, and a restriction endonuclease recognition site.

21. A binary oligonucleotide array according to claim 20, wherein said functional sequence is a priming region.

22. A binary oligonucleotide array according to claim 1, wherein each binary oligonucleotide is covalently linked to said surface through a long polymer chain.

23. A binary oligonucleotide according to claim 2, wherein said deoxyribonucleotides comprise at least one modified nucleotide.

24. A sectioned oligonucleotide array comprising an array of predetermined areas on a surface of a solid support, each area having therein, covalently linked to said surface multiple copies of an oligonucleotide, wherein said areas are physically separated from one another into sections, such that nucleic acids in an aqueous solution generated in one section cannot migrate to another section.

25. A sectioned oligonucleotide array according to claim 24 further comprising a lattice attached to said surface.

26. A sectioned oligonucleotide array according to claim 25, wherein said lattice is removably attached to said surface.

27. A sectioned oligonucleotide array according to claim 25, further comprising a cover removably attachable to said lattice.

28. A sectioned oligonucleotide array according to claim 24, wherein said sections comprise wells in said solid support.

29. A sectioned oligonucleotide array according to claim 28, further comprising a cover removably attachable to said solid support.

30. A sectioned oligonucleotide array according to claim 24, comprising a gel which physically separates said areas by preventing nucleic acids in an aqueous solution placed in one area from migrating to another area.

31. A sectioned oligonucleotide array according to claim 24, wherein said sections are mechanically separated from one another.

32. A sectioned oligonucleotide array according to claim 27, wherein said cover comprises a replica array.

33. A sectioned oligonucleotide array according to claim 29, wherein said cover comprises a replica array.

34. A sectioned array according to claim 24 wherein all of the oligonucleotides in individual areas are of the same sequence.

35. A sectioned array according to claim 24 wherein not all oligonucleotides in each area are of the same sequence.

36. A method of sorting a mixture of nucleic acid strands comprising the steps of:

a) providing a solution containing a mixture of nucleic acid strands in single-stranded form and

b) contacting said solution to a first binary oligonucleotide array of predetermined areas on a surface of a solid support, each area having therein, covalently linked to said surface, copies of a binary oligonucleotide, said binary oligonucleotide consisting of a constant nucleotide sequence adjacent to a variable nucleotide sequence, wherein the constant nucleotide sequence is the same for all oligonucleotides in the array, wherein said step of contacting is carried out under conditions

promoting perfect hybridization of said strands to said binary oligonucleotides.

37. A method according to claim 36 wherein said array is comprehensive.

38. A method according to claim 36 wherein said array is a 3' array.

39. A method according to claim 36 wherein said binary oligonucleotides are complementary to sequences that possibly occur in the strands in said mixture.

40. A method according to claim 39 wherein said array is comprehensive.

41. A method according to claim 36 wherein said array is a sectioned array, further comprising the step of amplifying strands hybridized in at least some of said areas to produce copies of said hybridized strands.

42. A method according to claim 36 further comprising removing strands that have not perfectly hybridized.

43. A method according to claim 42 further comprising adding a terminal extension to at least one terminus of the strands, said terminal extension having a sequence which substantially does not occur in the strands.

44. A method according to claim 43 wherein a terminal extension is added to the strands by ligation of hybridized strands to masking oligonucleotides, said masking oligonucleotides being also hybridized to said binary oligonucleotides.

45. A method according to claim 44 wherein a second terminal extension is added to the strands prior to said step of contacting, said second terminal extension being added to termini

not hybridized to said binary oligonucleotides during said step of contacting.

46. A method according to claim 42 further comprising releasing hybridized strands on a sectioned array into solution without mixing of material in said areas and rebinding them to said binary oligonucleotides followed by removing unhybridized strands.

47. A method according to claim 42 further comprising releasing hybridized strands in solution and rebinding to a replica array followed by removing unhybridized strands.

48. A method according to claim 42 wherein the mixture of nucleic acid strands comprises RNA.

49. A method according to claim 42 wherein the mixture of nucleic acid strands is comprised of DNA fragments obtained by site specific degradation.

50. A method according to claim 43 wherein the mixture is comprised of DNA fragments obtained by digestion with a restriction endonuclease and wherein the constant region of the binary oligonucleotide contains the complement of the restriction endonuclease recognition site, and wherein addition of the terminal extension restores the recognition site.

51. A method according to claim 42 further comprising generating complementary copies of hybridized strands.

52. A method according to claim 51 wherein the array is a 3' array wherein each binary oligonucleotide has its variable sequence adjacent to the 5' end of its constant sequence, and the copies are generated using a DNA polymerase and using the binary oligonucleotide as a primer.

53. A method according to claim 51 wherein the array is a 5' array wherein each binary oligonucleotide has its variable sequence adjacent to the 3' end of its constant sequence, and the copies are generated using a DNA polymerase using a primer hybridized to a 3' terminal extension of the hybridized strands, and the copies are then ligated to the 5' end of the binary oligonucleotides.

54. A method according to claim 44 further comprising amplifying the hybridized strands.

55. A method according to claim 51 further comprising removing the hybridized strands and amplifying the complementary copies of the hybridized strands.

56. A method according to claim 55 wherein the hybridized strands have 3' and 5' terminal extensions, and the amplification is a polymerase chain reaction.

57. A method according to claim 55 wherein the hybridized strands have a terminal extension and the amplification is linear.

58. A method according to claim 36 wherein said step of providing comprises digesting genomic DNA with a restriction endonuclease to create DNA fragments;

(a) modifying said fragments by adding a first constant sequence to their strands' 3' termini and a second constant sequence to their strands' 5' termini to create priming regions including restored restriction sites; and

(b) denaturing the modified fragments to form a mixture of single nucleic acid strands.

59. A method according to claim 58 wherein said array is a sectioned, comprehensive array, further comprising the step of amplifying strands hybridized in said areas by symmetric PCR.

60. A method according to claim 58 further comprising the step of amplifying said mixture of single nucleic acid strands by asymmetric PCR.

61. A method according to claim 36 wherein said binary oligonucleotides or portions thereof are complementary to terminal sequences that possibly occur in one end of the strands in said mixture and that are substantially non-complementary to internal sequences in the strands in said mixture.

62. A method according to claim 61 wherein said array is a sectioned array, further comprising the step of amplifying strands hybridized in at least some of said areas to produce amplified copies of said single nucleic acid strands.

63. A method according to claim 62 wherein said array is a comprehensive array.

64. A method according to claim 62 wherein said array is a 3' array.

65. A method according to claim 61 wherein said step of providing comprises digesting genomic DNA with a restriction endonuclease to create DNA fragments, modifying said fragments by adding a first constant sequence to their strands' 3' termini to create priming regions including restored restriction sites, and denaturing the modified fragments into a mixture of single nucleic acid strands.

66. A method according to claim 61 wherein said step of providing comprises digesting genomic DNA with a restriction endonuclease to create DNA fragments;

(a) modifying said fragments by adding a first constant segment to one of their strands' 3' and 5' termini to create priming regions including restored restriction sites; and

(b) denaturing the modified fragments into a mixture of denatured nucleic acid strands each having a priming region only at one end.

67. A method according to claim 66 wherein said first binary sorting array is a 3' array.

68. A method according to claim 67 further comprising the steps of

(a) generating an immobilized copy of each strand hybridized to the array by incubation with a DNA polymerase using the immobilized oligonucleotide as a primer and a hybridized strand as a template; and

(b) washing to remove from the array all materials not covalently bound to the array.

69. A method according to claim 68, wherein said step of modifying comprises adding a first constant sequence to their strands' 5' termini and wherein said 3' array contains binary oligonucleotides to which are hybridized masking oligonucleotides, further comprising the steps of

(a) ligating said masking oligonucleotides to denatured nucleic acid strands hybridized to said binary oligonucleotides such that their 3' termini are immediately adjacent to one of said masking oligonucleotides, and

(b) washing under conditions such that only strands so ligated will remain.

70. A method according to claim 69 wherein said step of adding a first constant sequence includes ligation of a double-stranded oligodeoxyribonucleotide adaptor.

71. A method according to claim 69 wherein said step of adding a first constant sequence includes ligation of a single-stranded oligoribonucleotide.

72. A method according to claim 68 wherein said step of modifying comprises adding a first constant sequence to their strands' 3' termini.

73. A method according to claim 72 wherein said first constant sequence is a homopolynucleotide tail added by extension of the strands' 3' termini by enzymatic extension.

74. A method according to claim 72 further comprising the step of adding a second constant sequence to the 3' termini of the immobilized copies.

75. A method according to claim 74 wherein said second constant sequence is a homopolynucleotide tail added by extension of said immobilized copies' 3' termini by enzymatic extension.

76. A method according to claim 68 wherein said first binary oligonucleotide array is a sectioned array, further comprising the step of amplifying said washed, immobilized copies to produce amplified copies.

77. A method according to claim 76 wherein said step of amplifying comprises PCR.

78. A method according to claim 76 wherein said first binary oligonucleotide array is a comprehensive array.

79. A method according to claim 76 further comprising contacting said amplified copies from at least one area of said 3' array to a second binary oligonucleotide array containing immobilized binary oligonucleotides whose constant sequence is identical or complementary to the 3' terminus of the immobilized copies.

80. A method according to claim 62 further comprising contacting said amplified copies from at least one area of said first binary oligonucleotide array to a second binary oligonucleotide array containing immobilized binary oligonucleotides that are com-

plementary to terminal sequences that possibly occur in either the other ends of said denatured nucleic acid strands or the complements of said other ends, and that are not complementary to internal sequences in the strands in said mixture or their complements.

81. A method according to claim 61 wherein said step of providing comprises digesting genomic DNA with a restriction endonuclease to create DNA fragments, and denaturing said fragments into a mixture of denatured nucleic acid strands.

82. A method according to claim 81 wherein said first binary oligonucleotide array is a 3' array containing binary oligonucleotides to which are hybridized masking oligonucleotides, further comprising the steps of ligating said masking oligonucleotides to denatured nucleic acid strands hybridized to said binary oligonucleotides such that their 3' termini are immediately adjacent to one of said masking oligonucleotides, washing under conditions such that only strands so ligated will remain, and generating an immobilized copy of each ligated strand by incubation with a DNA polymerase.

83. A method according to claim 82 further comprising the steps of adding a constant sequence to the 5' termini of the hybridized strands by ligation of a single-stranded oligoribonucleotide; incubating with a DNA polymerase to extend the immobilized copies; washing to remove from the array all materials not covalently bound to the array; and amplifying said washed, immobilized copies to produce amplified copies.

84. A method according to claim 83 wherein said step of amplifying comprises PCR.

85. A method according to claim 83 wherein said first sorting array is a comprehensive array.

86. A method according to claim 83 further comprising contacting said amplified copies from at least one area of said 3' array to a second binary array containing immobilized binary oligonucleotides whose constant sequence is identical or complementary to the 3' terminus of said immobilized copies.

87. A method according to claim 67 further comprising the steps of adding a constant sequence to the 3' termini of the immobilized copies by enzymatic extension thereof; washing to remove from the array all materials not covalently bound to the array; and amplifying said washed, immobilized copies to produce amplified copies.

88. A method according to claim 87 wherein said step of amplifying comprises PCR.

89. A method according to claim 87 wherein said first sorting array is a comprehensive array.

90. A method according to claim 87 further comprising contacting said amplified copies from at least one area of said 3' array to a second terminal binary array containing immobilized binary oligonucleotides whose constant sequence is identical or complementary to the 3' terminus of said immobilized copies.

91. A method according to claim 61 wherein said step of providing comprises digesting genomic DNA with a site-specific cleaving agent to create DNA fragments.

92. A method according to claim 91 wherein said agent is an endonuclease.

93. A method according to claim 91 wherein said agent is a chemical agent.

94. A method according to claim 61 wherein said nucleic acid strands are cDNA strands.

95. A method according to claim 61 wherein said nucleic acid strands are RNA strands.

96. A method according to claim 95 wherein said RNA strands are eukaryotic mRNA strands, and wherein said step of providing comprises removing 5'-cap structures.

97. A method according to claim 95 wherein said RNA strands lack a poly(A) tail.

98. A method according to claim 61 wherein said step of providing comprises digesting genomic DNA with a restriction endonuclease to create DNA fragments;

(a) modifying said fragments by adding a first constant sequence to their strands' 3' termini and a second constant sequence to their strands' 5' termini to create priming regions including restored restriction sites; and

(b) denaturing the modified fragments into a mixture of single nucleic acid strands.

99. A method according to claim 98 wherein the 3' priming regions are complementary to the 5' priming regions.

100. A method according to claim 99 wherein said array is a 3' array, further comprising the steps of

(a) generating an immobilized copy of each strand hybridized to the array by incubation with a DNA polymerase; and

(b) washing to remove from the array all materials not covalently bound to the array.

101. A method according to claim 100 wherein said array is a sectioned array, further comprising the step of amplifying strands hybridized in at least some areas by PCR to produce amplified copies of each said immobilized copy.

102. A method according to claim 101 wherein said array is a comprehensive array.

103. A method according to claim 99 wherein addition of said first constant sequence and said second constant sequence includes ligation of a double-stranded oligodeoxyribonucleotide adaptor to the strands' 5' termini.

104. A method according to claim 99 wherein addition of said first constant sequence and said second constant sequence includes ligation of a single-stranded oligonucleotide to the strands' 5' termini.

105. A method according to claim 99 wherein addition of said first constant sequence and said second constant sequence includes enzymatic extension of the strands' 3' termini by the synthesis of a homopolynucleotide tail.

106. A method according to claim 101 further comprising contacting said amplified copies from at least one areas of said 3' array to a second binary array under conditions promoting hybridization of said amplified copies to the binary oligonucleotides in said second array.

107. A method according to claim 106 wherein said amplified copies are produced by symmetric PCR and wherein said second array is a 3' array.

108. A method according to claim 106 wherein said first array and said second array are comprehensive.

109. The product of a method according to claim 100.

110. A method of sorting a mixture of nucleic acid strands comprising the steps of

a) providing a solution containing a mixture of nucleic acid strands in single stranded form, and

b) contacting said solution to an oligonucleotide array of predetermined areas on a surface of a solid support, each area having therein copies of an immobilized oligonucleotide, the nucleotide sequence of immobilized oligonucleotides in separate areas being different, wherein said contacting is performed under conditions that promote the formation of perfect hybrids.

111. A method according to claim 110 wherein said array is comprehensive.

112. A method according to claim 110 wherein the array is sectioned.

113. A method according to claim 110 wherein the immobilized oligonucleotides are between 6 and 30 nucleotides long.

114. A method according to claim 110 wherein the array is a 3' array.

115. A method according to claim 110 wherein the array is a 5' array.

116. In a method wherein two nucleic acid strands are ligated to each other in order to form a recombinant product, the improvement comprising hybridizing first nucleic acid strands to immobilized oligonucleotides in an oligonucleotide array prior to ligation to second nucleic acid strands, said oligonucleotide array comprising an array of predetermined areas on a surface of a solid support, each area having copies of an oligonucleotide immobilized thereon.

117. A method according to claim 116 wherein the first nucleic acid strands have different nucleotide sequences in each area of the array.

118. A method according to claim 116 wherein the second nucleic acid strands have different nucleotide sequences in each area of the array.

119. A method according to claim 116 wherein the array is a comprehensive array.

120. A method according to claim 116 wherein the oligonucleotides immobilized in each area are of the same length.

121. A method according to claim 116 wherein the oligonucleotides consist of the group consisting of deoxyribonucleotides, ribonucleotides, mixed deoxyribonucleotides and ribonucleotides, modified deoxyribonucleotides, modified ribonucleotides, and non-standard nucleotides.

122. A method according to claim 116 wherein the second nucleic acid strands are not also hybridized to the immobilized oligonucleotides.

123. A method according to claim 122 wherein the second nucleic acid strands are strands of double stranded nucleic acids.

124. A method according to claim 123 wherein the set of double stranded nucleic acids has one end adapted for ligation to blunt ends formed by hybridization of the first nucleic acids to the immobilized oligonucleotides.

125. A method according to claim 116 wherein non-ligating termini of the first nucleic acid strands and the double stranded nucleic acids contain priming regions for amplification.

126. A method according to claim 125 wherein following ligation of the first nucleic acids to the second nucleic acids, polymerase chain reaction amplification is carried out.

127. A method according to claim 124 wherein the double stranded nucleic acids are ligated to the immobilized oligonucleotide using RNA ligase prior to ligation of the first nucleic acid strands and the second nucleic acid strands.

128. A method according to claim 123 wherein the second set of nucleic acids is the same in every area of array.

129. A method according to claim 123 wherein the first nucleic strands are hybridized to the immobilized oligonucleotides while contained in a mixture of one or more different strands, said different strands having terminal sequences different from corresponding termini to be ligated of the first nucleic acid strands.

130. A method according to claim 116 wherein both the first nucleic acid strands and the second nucleic acid strands are hybridized to the immobilized oligonucleotides in the array prior to ligation.

131. A method according to claim 130 wherein both the first and second nucleic acid strands contain priming regions at their non-ligating termini.

132. A method according to claim 131 wherein the first and second nucleic acid strands are amplified in a polymerase chain reaction following ligation.

133. A method according to claim 130 wherein both the first and second nucleic acids are, prior to hybridization to the immobilized oligonucleotides, contained in mixtures of nucleic acids having terminal sequences different from the corresponding termini to be ligated of the first nucleic acid strands and the second nucleic acid strands.

134. A method according to claim 36 further comprising sorting the hybridized nucleic acid strands or their copies in an area of

the first binary array by contacting them to a second oligonucleotide array.

135. A method according to claim 134 wherein the strands or their copies are contacted to all areas of the array.

136. A method according to claim 36 wherein the nucleic acid strands are contacted to all areas of a second binary array.

137. A method according to claim 134 wherein cleavable primers are used following said step of contacting for amplification of hybridized strands.

138. A method according to claim 137 further comprising cleaving the cleavable primers from the strands and adding new terminal extensions.

139. A method according to claim 134 wherein the contents of an area of the first binary array are contacted with only predetermined areas of a second binary array.

140. A method according to claim 36 further wherein contents in an area of the binary array are contacted with the corresponding area of a replica array.

141. A method according to claim 134 wherein the second oligonucleotide array is a second binary array.

142. A method for introducing a site directed mutation into a nucleic acid strand on an oligonucleotide array using a partial, said partial corresponding to a region of the nucleic acid strand adjacent to the location of the site directed mutation to be introduced, comprising the steps:

(a) separately ligating said partial to the free terminus of a preselected immobilized oligonucleotide in the oligonucleotide array to obtain a mutated partial, said oligonucleotide array comprising an array of predetermined areas on the

surface of a solid support, each area having therein a preselected immobilized oligonucleotide, said preselected oligonucleotide having a sequence adapted to introduce a mutation to the partial added to the area; and

(b) generating, using the mutated partial, a nucleic acid containing the mutation.

143. A method according to claim 142 wherein step b is accomplished by

(a) hybridizing a complementary copy of the mutated partial to a template having the complementary sequence of the terminal portion of the nucleic acid strand which is not contained in the partial; and

(b) carrying out a polymerase reaction, a ligation reaction or both a polymerase reaction and ligation reaction to join the remaining region of the nucleic acid strand to the mutated partial.

144. A method for making immobilized partial copies of a nucleic acid strand on a 3' or 5' oligonucleotide array, comprising the steps:

(a) hybridizing the strand to the array by an oligonucleotide segment contained in the strand, said array comprising predetermined areas on a surface of a solid support, each area having therein immobilized oligonucleotides consisting of a predetermined variable sequence, said hybridization taking place under conditions that promote the formation of perfect hybrids of the length of the immobilized oligonucleotide in each area, and

(b) where the strand is hybridized to a 3' array, enzymatically extending the immobilized oligonucleotide using the hybridized strand as a template, and where the strand is hybridized to a 5' array, hybridizing a primer to a priming region contained in the 3' terminus of the hybridized strand, then enzymatically extending the primer to form an extension product, then ligating the extension product to the immobilized oligonucleotide.

145. A method according to claim 144 wherein the strand is hybridized to a 3' array, further comprising amplifying the immobilized partial copies using a primer or promoter complement appropriate to hybridize to a priming region or promoter sequence at the immobilized partial copies' 3' termini, and an appropriate polymerase.

146. A method according to claim 144 wherein the oligonucleotide array is substantially comprehensive.

147. A method according to claim 146 wherein a substantially complete set of immobilized partial copies is generated on the array by

(a) hybridizing the strand to the array by substantially all oligonucleotides present in the strand;

(b) performing step (b) on all hybridized strands.

148. A method according to claim 146 wherein a substantially complete set of amplified partials is generated on a 3' array by

(a) hybridizing the strand to the 3' array by substantially all oligonucleotides present in the strand;

(b) performing step (b) on all hybridized strands; and

(c) amplifying substantially all immobilized partial copies by using a primer or promoter complement appropriate to hybridize to a priming region or promoter sequence at the partial copy's fixed terminus, and an appropriate polymerase.

149. A method according to claim 148 wherein following step (a) unhybridized and imperfectly hybridized strand copies are removed.

150. A method according to claim 149 wherein the array is sectioned.

151. A method according to claim 150 wherein the strand is contained in a mixture of strands which are subjected to the same steps on the array.

152. A method according to claim 151 wherein the priming region is a terminal extension introduced in all strands in the mixture.

153. A method according to claim 149 wherein the priming region or promoter is added to the 5' terminus of the nucleic acid strand prior to hybridizing the strand to the array.

154. A method according to claim 150 further wherein the oligonucleotide content in an area of the array is surveyed.

155. The product of a method according to claim 144.

156. The product of a method according to claim 146.

157. A method according to claim 144 wherein the strand is contained in a mixture of sorted strands subjected to the method, said mixture of sorted strands being from an area of a sorting array.

158. A method according to claim 157 further wherein mixtures of strands from different areas of the sorting oligonucleotide array are hybridized to the 3' or 5' oligonucleotide array.

159. A method according to claim 144 wherein the nucleic acid is a previously prepared partial.

160. A method according to claim 145 further comprising sorting partials or their copies from an area of the oligonucleotide array on a second oligonucleotide array.

161. A method according to claim 145 further comprising sorting partials or their copies from an area of the oligonucleotide array according to variable sequences adjacent their fixed ends on a binary oligonucleotide array.

162. A method of claim 144 further comprising ligating a partial or its copy in single stranded or double stranded form to a second nucleic acid strand.

163. A method according to claim 162 wherein the second nucleic acid strand is a previously obtained partial.

164. A method according to claim 145 further wherein a cleavable primer, at an end of a partial to be ligated, is used for amplification, and further comprising cleaving the primer and then ligating the partial to a second nucleic acid strand.

165. A method according to claim 162 further comprising exponentially amplifying ligated product using priming regions at non-ligated termini.

166. A method according to claim 165 further wherein the priming regions at the non-ligated termini of the ligated product are adapted to permit amplification only of the ligated product.

167. A method according to claim 144 further wherein a partial obtained is ligated to an oligonucleotide or to a second nucleic acid strand adapted to introduce a site directed mutation, with respect to the nucleic acid strand that the partial was generated from, at the ligated terminus of the partial.

168. A method according to claim 167 wherein the oligonucleotide is immobilized in a second oligonucleotide array.

169. A method for sorting partials by their variable termini on a binary oligonucleotide array, which partials have been prepared by random chemical or enzymatic degradation of one or more nucleic acid strands, said binary array comprising an array of predetermined areas on a surface of a solid support, each area having therein copies of a binary oligonucleotide of a predetermined sequence, said binary oligonucleotide consisting of a constant nucleotide sequence adjacent to a variable nucleotide

sequence, said variable nucleotide sequence being at the free end of the binary oligonucleotides, said binary oligonucleotide also having a complementary masking oligonucleotide hybridized to all or a part of the constant nucleotide sequence, including the portion of the constant nucleotide sequence adjacent the variable nucleotide sequence, comprising the steps of:

(a) hybridizing the partials to the array by their termini under conditions that promote the formation of perfect hybrids; and

(b) ligating the termini of the partials to the masking oligonucleotide.

170. A method for obtaining information for determining the sequence of a nucleic acid strand comprising

(a) generating a substantially complete set of partials of the nucleic acid strand; and

(b) for groups of partials, having the same terminal variable nucleotide sequence of predetermined length, separately determining the presence and sequence of all variable oligonucleotides of the predetermined length.

171. In a method for surveying oligonucleotide content of a nucleic acid strand as part of a sequencing method wherein the strand is hybridized to a comprehensive oligonucleotide array, and the presence of hybridized strands in areas of the array is detected, the improvement comprising:

(a) preparing a substantially complete set of partials of the strand prior to surveying;

(b) sorting the partials by their variable ends on an oligonucleotide array, and

(c) separately surveying oligonucleotide content of each group of sorted partials.

172. A method according to claim 171 wherein the strand is in a mixture of strands which are subjected to the same steps.